

## ***BIG DATA: MODOS DE FAZER, COMPARAR E GOVERNAR***

Sara R. Munhoz<sup>1</sup> 

Vanessa Parreira Perin<sup>2</sup> 

Magda dos Santos Ribeiro<sup>3</sup> 

<sup>1</sup>Universidade Estadual de Campinas,  
Departamento de Política Científica e Tecnológica. Campinas, São Paulo, Brasil.

<sup>2</sup>Universidade Federal de São Carlos,  
Programa de Pós-Graduação em Antropologia Social. São Carlos, São Paulo, Brasil.

<sup>3</sup>Universidade Federal de Minas Gerais (UFMG),  
Departamento de Antropologia e Arqueologia. Belo Horizonte, Minas Gerais, Brasil.

E-mails: sararmunhoz@gmail.com;  
vanessa\_pperin@hotmail.com.br; magdaribeiro@ufmg.br

Em novembro de 2022, a empresa OpenAI lançou e disponibilizou a qualquer usuário cadastrado o software batizado como Chat GPT. Uma celebridade maquínica instantânea. Ainda mais estrondoso do que muitas outras ferramentas, plataformas e aplicativos que o antecederam, o nascimento do chatbot foi descrito como uma tecnologia tão revolucionária quanto o surgimento da própria internet. Uma Nova Era. A imagem não é isolada nem aleatória. Uma marcação de tempo que se destaca abruptamente da que a antecedeu e dá início a uma ordem renovada das coisas. As Eras (da Informação, Digital, dos Dados, das Inteligências Artificiais etc.) têm ora se substituído, ora se sobreposto em um sistema de crescente aceleração do *kronós*, cujos efeitos, tal como já vislumbrados por Paul Virilio (1977), só podem ser apreensíveis por uma dromologia, uma ciência atenta à velocidade.

Com o advento e a rápida difusão do GPT, órgãos públicos e privados imediatamente passaram a conjecturar e a experimentar os usos possíveis e ainda latentes para a ferramenta em suas atividades cotidianas, reordenando por completo o cenário das *potencialidades* mercadológicas e de conhecimento, para aproveitarmos um conceito mobilizado por Vanessa Perin neste dossiê. Ao mesmo tempo, os perigos iminentes aos direitos constitucionalmente garantidos rapidamente começaram a ser apontados. E os alertas têm sido abundantes. As próprias discussões legislativas, que entre 2021 e 2022 engatinhavam no Brasil rumo

a propostas de regulamentação das plataformas e dos usos da IA, depararam-se com um cenário significativamente transformado, que exigiria a inclusão de muitas novas variáveis nas tentativas de circunscrição normativa do legal e do ilegal no *continuum* on-line-off-line contemporâneo. O recente e intenso avanço no desenvolvimento das tecnologias baseadas na internet, como plataformas de redes sociais e aplicativos diversos para dispositivos móveis, também fez com que muitos líderes do setor privado se encontrassem despreparados para lidar com o enorme volume de dados gerados em altíssima velocidade — como demonstra a recente crise de gestão na mencionada OpenAI.<sup>1</sup> Ademais, se convencionalmente novas tecnologias eram desenvolvidas e testadas em laboratórios, sendo posteriormente apresentadas ao público em veículos de mídia para que só então passassem a ser adotadas de modo mais amplo, o rápido desenvolvimento e a pronta aceitação destes artefatos pela comunidade de usuários têm limitado significativamente o tempo de maturação e reflexão crítica sobre seus efeitos a partir de uma perspectiva acadêmica (Alabdullah et al. 2018; Gandomi & Haider 2015).

Este dossiê se insere no amplo espectro dos estudos interdisciplinares acerca da emergência de novas tecnologias, com ênfase nos sistemas digitais e informacionais que envolvem o uso de dados. Quando proposto à *Mana*, estávamos justamente à época da ruptura estrondosa gerada pela popularização das inteligências artificiais generativas. Mas ele vem à luz quando a convivência com elas já começa a se assentar como uma parte corriqueira, ordinária de nossas vidas. Seu léxico já não nos surpreende, seus alcances já não nos estupefazem como o faziam há dois anos, ainda que a cada dia novos usos e possibilidades nos sejam apresentados como pontos de inflexão extraordinários. Temos sido rapidamente transformados em sujeitos dessa nova Era. Nossos corpos, sentidos e atenções, nossos dedos e retinas, respiração e massa cerebral têm se familiarizado com a velocidade que os transforma radicalmente. Experimentamos a materialização palpável de processos não inéditos, tampouco isolados, mas ainda assim colossais de inovação e aniquilação, se acompanharmos o argumento de Magda Ribeiro e Manuel Luz, também neste dossiê.

O conjunto de trabalhos aqui reunidos dedica-se aos dados. Aos sistemas e às práticas que os reúnem, relacionam e transformam. Tratamos daqueles que apenas quando colecionados em grandes volumes, suficientemente despersonalizados e satisfatoriamente postos em contato em aparatos infraestruturais que os organizam, podem ser, potencial e alegadamente, transformados em recursos infindáveis. *Big data*. Ainda que as etnografias aqui selecionadas não descrevam as inteligências maquinicas generativas, embora, enfim, se dediquem a uma Era que tem sido descrita como em vias de superação, os trabalhos escancaram o aspecto absolutamente criativo e criador destes que têm se tornado o átomo fundamental de nossa imaginação, política e epistemologia: os dados. Veremos

como eles incitam desejos, fincam terra, firmam cooperações internacionais, delimitam territórios, abrem mercados, circunscrevem a democracia e a justiça. Nunca isolados, mas inseridos em complexos emaranhados sociotécnicos, que os aproximam entre si, mas também os conectam a pessoas, ferramentas, e instituições — engajam modalidades de governos, de controle, de hierarquias, submissões e padronizações. Postos em contato e em comparação, ordenam o mundo que se empenham em criar.

A esta introdução caberá apresentar brevemente o que, afinal, dizemos quando falamos em *big data*, os diálogos entre as etnografias aqui reunidas e sua inserção na produção nacional e internacional a respeito deste tema. Mas antes disso, gostaríamos de desacelerar um instante e destacar um aspecto que julgamos fundamental nesta publicação conjunta: em meio ao *frenesi* dos excessos e das transformações alegadamente irrefreáveis, que afetam inclusive os modos de organizarmos nosso pensamento, nossa escrita e nossas publicações, é importante mencionar que este dossiê foi gestado em outro compasso: o das pausas, dos respiros, do trabalho artesanal. Alguns de seus artigos apresentam desdobramentos de pesquisas de longa duração, como teses de doutorado. Outros envolvem experimentações etnográficas, teóricas e metodológicas que demandaram tempo considerável de maturação e, mais do que isso, significativa hesitação (Stengers 2007). O resultado, esperamos, se coloca como resistência firme — ainda que não ingênua — ao culto do imediatismo. Não nos encabulamos ao apresentar dados de pesquisa que, tão logo publicados, já escancaram sua obsolescência — o Twitter já não existe tal como há meses atrás, o Projeto Embrapa-Moçambique foi encerrado, os procedimentos de filtragens dos documentos no STJ também se transformaram significativamente, as técnicas de manipulação de organismos genéticos aqui descritas já foram superadas etc. Isto porque defendemos que a minúcia e a densidade que caracterizam as descrições etnográficas, e o contato entre campos de pesquisa tão distintos, podem fazer com que emergjam, uma vez mais e de modo renovado, reflexões que afetam muito diretamente temas centrais e perenes à atenção antropológica, como os muitos modos de se fazer/criar, comparar e governar.

## 1

De origem incerta, os primeiros usos da expressão *big data* têm sido localizados nas conversas em mesas de almoço da Silicon Graphic Inc. (empresa norte-americana de soluções para computação de alta performance) no início da década de 1990, posteriormente disseminados para outros espaços por meio de empresas que buscavam consolidar um mercado para os produtos baseados em uma nova

escala de coleta, armazenamento e processamento de dados (Alabdullah et al. 2018; Lohr 2013). Segundo Francis X. Diebold, economista da Universidade da Pensilvânia que publicou um dos primeiros estudos acadêmicos utilizando a expressão *Big Data* (em maiúsculas), as origens do termo são intrigantes e um tanto obscuras, envolvendo não apenas a indústria e a ciência da computação, mas, fundamentalmente, a estatística e a econometria (Diebold 2019). A economia como terreno fértil para a emergência do termo também é apontada por Steve Lohr (2013), colunista do *The New York Times*, que se embrenhou na busca pela história etimológica do conceito, entendendo-o não apenas como uma grande quantidade de dados, mas como diferentes tipos de dados tratados de novas maneiras.

Neste sentido, Lohr destaca que alguns dos métodos estatísticos e algorítmicos agora incluídos no kit de ferramentas de *big data* têm a sua herança na modelização econômica e no mercado financeiro. Embora reconheça o pioneirismo acadêmico de Diebold, o jornalista situa os primórdios do uso conjugado dos termos *big* e *data* nas palestras conduzidas pelo cientista da computação e empresário John Mashey, quando este buscava tanto explicar o conceito quanto apresentar os produtos da Silicon Graphics Inc. (empresa na qual era vice-presidente e cientista chefe). A popularidade atual do conceito também tem sido atribuída às iniciativas promocionais da IBM e de outras empresas líderes de tecnologia, que buscaram investir na construção de um nicho de mercado para a análise de dados (Gandomi & Haider 2015). Vale lembrar que décadas antes, coadunados com os esforços de guerra que marcaram o século XX, os estudos da cibernética já se enveredavam por pesquisas que resultaram tanto no computador analógico quanto em novas tecnologias digitais, além de terem contribuído para direcionar a teoria econômica para a hegemonia neoliberal contemporânea (Cesarino 2021; Mirowski & Nik-Khah 2017).

Se a origem da expressão *big data* é controversa, seu significado não é menos disputável, uma vez que nenhum termo captura completamente seu referente (Boellstorff & Maurer 2015). Como aponta Manovich (2011a), é na indústria da informática que encontramos uma definição mais circunscrita: *big data* refere-se aos conjuntos de dados cujo tamanho está além da capacidade de captura, gerenciamento e processamento dos *softwares* comuns, requerendo "supermáquinas" para serem tratados. O tamanho dos conjuntos de dados considerados *big*, porém, tem sido alvo de constante transformação, situando-se atualmente na casa das dezenas de *terabytes* e de múltiplos *petabytes*. Além disso, muito do que antes se considerava um trabalho maquínico para supercomputadores pode agora ser analisado em *desktops* padrões (boyd & Crawford 2012). De modo geral, grandes volumes de dados digitais, produzidos em escala industrial, têm sido frequentemente alocados na categoria *big data*. Todavia, uma estrutura comum para

qualificar este tipo de conjunto de dados, para além de suas quantidades, têm sido as três dimensões definidas por Laney (2001) como fundamentais para a gestão dos dados de e-comércio: volume, velocidade e variedade. A partir destes elementos, muitos outros traços foram sendo agregados como possíveis qualificadores dos *big data*: veracidade, variabilidade, versatilidade, volatilidade, vitalidade, valor – e também exaustividade, indexicalidade, extensionalidade, escalabilidade, refino (Kitchin & McArdle 2016). A lista é extensa. Em uma perspectiva crítica, Lupton (2015) sugere, ainda, um conjunto de palavras com “p” a serem associadas aos *big data*: perverso, pessoal, produtivo, parcial, prático, preditivo, político, provocativo, privado, polivalente, polimorfo.

Neste dossiê, entretanto, não nos voltamos para a construção de narrativas de origem ou para o estabelecimento de definições fechadas. Antes, nos aproximamos da abordagem de Boellstorff (2013), que ao tratar *big data* como substantivo comum ressalta sua constante possibilidade de reconfiguração. Sem desconsiderar as importantes transformações que o conceito sofreu ao longo das últimas três décadas, bem como, e principalmente, os modos muito particulares pelos quais é definido localmente, os artigos que compõem o dossiê investem em uma abordagem que enfatiza, analítica e descritivamente, o potencial de hiper-relacionalidade autorizado pela comparação em larga escala e pelas referências cruzadas sempre abertas a novos arranjos (Walford 2017, 2018). Como apontam Gandomini e Haider (2015), a relatividade dos grandes dados, isto é, o fato de que as múltiplas qualificações mencionadas acima são interdependentes e modificam-se umas em relação às outras, pode ser um denominador comum para pensarmos os *big data*. É neste sentido que muitos dos artigos do dossiê mobilizam boyd e Crawford (2012), na medida em que os abordam como um fenômeno sociotécnico que tem menos a ver com o volume do que com a capacidade de construir referências cruzadas entre massivas quantidades de informação, que precisam ser interrogadas criticamente. Eles não apresentam apenas um valor em si, mas também – e principalmente – em seu potencial de transformação revelado por técnicas comparativas sempre renovadas (Walford 2021). De modo semelhante, Boellstorff e Maurer (2015) nos lembram que os sistemas computacionais móveis e digitais, que permitem a geração, a coleta e a análise em larga escala de informações sobre as atividades, as localizações e as interações entre pessoas e dispositivos, são sempre formados por relações que se estendem para além dos dados em si. O que conta como dado (assim como seus referentes) é definido por processos sociopolíticos, dado este que se transforma em “tempo real” e atravessa as noções de natureza e cultura.

Se, junto a Strathern (2017), consideramos que um dos pressupostos do conhecimento antropológico é tratar de relações entre relações, atentar à hiper-relacionalidade dos dados digitais pode adensar e estender nossos empreendimentos comparativos. Mas, neste ponto, é importante enfatizarmos que essa aposta

analítica na hiper-relacionalidade e nas *potencialidades* ilimitadas dos dados decorre, antes de mais nada, da observação etnográfica. Não se trata de ferramenta conceitual escolhida de antemão que se sobrepõe e enrijece o material de nossas pesquisas. Funcionários públicos lotados em setores administrativos de tribunais, cientistas imersos na floresta amazônica, políticos e empresários interessados em cooperações internacionais, plataformas digitais transnacionais e seus usuários, empresas de tecnologia agrícola investindo na produção do interior do país, economistas acadêmicos, pesquisadores especialistas em biotecnologia: cada um à sua maneira, os interlocutores dessas pesquisas justificam suas ações, suas ambições e inseguranças nos "devaneios de um mundo conectado", como o chamou Werner Herzog (*Lo and Behold* 2016). E, assim, nos auxiliam a compreender os *big data* como um fenômeno cultural, tecnológico e de produção de conhecimento que se baseia na interação entre tecnologias e enunciados utópicos e distópicos (boyd & Crawford 2012). Ademais, nos alinhamos à proposta da coletânea organizada Knox e Nafus (2018) que, ao reunirem trabalhos etnográficos produzidos em um mundo saturado por dados, buscam romper com o "hype" da mídia e dos negócios financeiros-empresariais sobre os *big data* e suas promessas. Em alternativa, os autores propõem não apenas agregar dados empíricos sobre realidades investigadas, mas encarar o empreendimento etnográfico – com sua estratégia descritiva e analítica características – como um caminho fecundo para gerar novas formas de compreensão e teorização dos dados digitais. Abrindo, ainda, a possibilidade de renovar problemas caros à própria teoria antropológica a partir dessas etnografias (Cesarino 2021), como veremos mais detidamente na seção a seguir.

No que concerne ao problema da escalabilidade dos dados, bem como de sua capacidade de expandir e multiplicar, temos observado, por meio de diferentes pesquisas etnográficas, a relação entre dados e a noção recorrente de que, por meio de sua obtenção e acesso, a "natureza foi conquistada" – um trunfo do *design* de precisão, não apenas na computação, mas também nos negócios empresariais (Rabinow 1996; Tsing 2019). Tal escalabilidade permite criar mundos aproximáveis tanto pela semelhança como pelo contraste, abrindo novas questões para empreendimentos comparativos realizados pela antropologia (Strathern 2017, 2021; Seaver 2018; Douglas-Jones; Walford & Seaver 2021): como conduzir práticas relacionais entre elementos que escapam às padronizações dos dados ou que permitem apreender sua forma de operar? Qual o lugar da diferença, do dessemelhante, do particular, nas composições produzidas pelas grandes coleções de dados? Somadas a estas questões meta-antropológicas fundamentais, as etnografias reunidas neste dossiê mobilizam-se, então, por ainda outros três problemas condutores, que a elas se articulam e com elas se emaranham. O que os dados, quando suficientemente volumosos para serem considerados *big*, são capazes de fazer ou de criar? De que artifícios e artefatos lançam mão para que

esses seus processos criativos possam se efetivar? E, finalmente, como operam para que essas próprias criações se apresentem como elementos suficientemente legíveis, domesticados, governáveis?

Neste sentido, todas as publicações aqui reunidas voltam-se a práticas técnico-científicas e mercadológicas interessadas nos arsenais abrangentes e heterogêneos de coleta, mensuração, análise e divulgação dos *big data*, discutindo, cada uma a seu modo, as vinculações entre volumosos bancos de dados e aparatos diversos – aplicativos, satélites, microscópios, documentos, algoritmos, máquinas agrícolas – capazes de arranjar e distribuir elementos informacionais por meio da prática comparativa. As etnografias examinam tanto a materialidade destes encontros quanto a força produtiva da imaterialidade neles implicadas, em práticas como as de digitalização; de coleta, armazenamento e processamento de *big data* em diferentes setores e instituições; de pulverização das linguagens algorítmicas e estatísticas; de exigências crescentes por transparência, eficiência, celeridade e otimização; de demandas por desenvolvimento, sustentabilidade e produtividade incontornavelmente vinculadas a estes dados. Os trabalhos enfrentam as transformações éticas e epistemológicas relacionadas aos *big data*, bem como os efeitos da mobilização de conhecimentos computacionais pelos especialistas nesses variados ambientes. Descrevem, enfim, como técnicas bastante distintas entre si – tais como a jurídica, a burocrática, a cartográfica, a agrícola, a econômica, a financeira, a científica – deslocam-se por diferentes escalas vinculando-se ao ecossistema dos dados e investigam o que, nestes movimentos, são capazes de produzir.

## 2

A produção antropológica interessada nos computadores, em sua lógica, em sua linguagem e em seus instrumentais é antiga. Sabe-se da vinculação estreita da antropologia com os primeiros desenvolvimentos da cibernética (Ruyer 1954; Wiener 2017), manifesta principalmente na participação de Margareth Mead e Gregory Bateson nas revolucionárias *Macy Conferences* entre os anos 1940 e 1950 (Bateson 1972; Cesarino 2022). Mas também são conhecidos os desejos estruturalistas de mobilização das máquinas de computar nos estudos dos mitos e do parentesco, por exemplo (Hymes 1965; Seaver 2014; Pedersen 2023).

Com a popularização dos computadores pessoais, teve início a consolidação de um novo campo de estudos, voltado à especificidade das relações sociais estabelecidas no mundo on-line, à definição do virtual e aos elementos caracterizadores da cibercultura (Escobar 1994; Lévy 2003, 2010; Rifiotis et al. 2010). Destacamos, neste sentido, o pioneirismo do GrupCiber (Grupo de

Pesquisa em Ciberantropologia), que desde meados da década de 1990 vem fomentando debates fundamentais para a consolidação desse campo de estudos na antropologia brasileira (Rifiotis 2016; Segata et al. 2012). Sempre a partir de uma chave etnográfica e em diálogo com a Teoria Ator-Rede, as pesquisas do grupo têm se voltado às redes sociotécnicas "onde pessoas, artefatos e outros seres são cruzados e coproduzidos com e pelas tecnologias digitais" (Segata & Rifiotis 2016:10).

A atenção antropológica, aos poucos, passou a reivindicar a atualização das reflexões a respeito da coleta de seus dados, do trabalho de campo e do método etnográfico. Nos anos que se seguiram, as etnografias interessadas no digital começaram a se desdobrar em várias frentes e em diferentes abordagens teórico-metodológicas. Entre o final dos anos 1990 e o início dos anos 2000, a publicação de importantes etnografias e manuais consolidaram os estudos sociais da internet e a antropologia digital (Miller & Slater 2000; Miller 2001; Hine 2005, 2015; Boellstorff & Maurer 2015; Kozinets 2010; Horst & Miller 2012). Valerá mencionar que a dissertação de mestrado de Carolina Parreiras (2008), autora que integra este dossiê, é um dos trabalhos precursores desse novo campo que começava a se consolidar no Brasil. Aos poucos, esses novos trabalhos passaram a enfatizar o *continuum* on-line-off-line e as políticas etnográficas para o campo da cibercultura, enfrentando as críticas iniciais de que as pessoas estariam ausentes do ciberespaço e de que não seria possível produzir etnografias no ambiente on-line ou mediadas por plataformas e outros artefatos digitais (Segata & Rifiotis 2016).

Paralelamente, e extrapolando os limites disciplinares da antropologia, multiplicou-se a bibliografia a respeito do que se chamou de culturas de mediação eletrônica (Eisenlohr 2004), estudos dos mundos virtuais (Malaby 2009), das mídias digitais (Coleman 2010), ou ainda as ciências sociais computacionais (Valentim & Pavesi 2021). Trabalhos de pioneiros como Friedrich Kittler (2017), Katherine Hayles (1999) e Alexander Galloway (2004) ganharam novas leituras e investimentos nos chamados *hardware studies*, com o apelo de que cientistas sociais interessados nos computadores, na internet e nas formas de comunicação eletronicamente mediadas voltem-se à sua materialidade, ao funcionamento de suas infraestruturas, à lógica e à arquitetura das máquinas propriamente ditas (Ferreira 2006; Vismann & Krajewki 2007).<sup>2</sup> Por outro lado, Lev Manovich (2002, 2003) e Matthew Fuller (2008) encabeçaram os chamados *software studies*, que se apresentavam como um complemento necessário aos já mais bem consolidados *media studies*. Essas abordagens, que encaram o "*software* como cultura" (Fuller 2006) — não como simples tecnologia, mas como um novo *medium* por meio do qual podemos pensar e imaginar diferentemente — ganharam muita força a partir da segunda década do século XXI. O *software*, argumenta Manovich, é uma "camada que permeia todas as áreas das sociedades contemporâneas" e, portanto, as análises



a respeito das "técnicas de controle, comunicação, representação, simulação, análise, tomada de decisão, memória, visão, escrita e interação" não podem ser "completas" sem que ele seja efetivamente incluído (Manovich 2013:15, tradução nossa). Neste sentido, as experiências técnicas em linguagem de programação e em escrita e ensino de *software* poderiam despontar novas e enriquecedoras metodologias, inclusive entre pesquisas propriamente antropológicas.<sup>3</sup>

Ainda antes da pandemia da COVID-19, as publicações de Horst e Miller (2012), Geismar e Knox (2012), Pink et al. (2016), Knox e Nafus (2018) foram importantes guias a respeito de como conduzirmos etnografias em/dos/sobre contextos digitais (Ramos & Freitas 2017; Cesarino 2021). Mas trabalhos explicitamente dedicados às transformações no método etnográfico depois da popularização da internet e das redes sociais proliferaram significativamente nos anos de isolamento social, como resposta às exigências de que antropóloga(o)s reformulassem suas atividades de pesquisa mesmo quando o objeto de investigação não envolvia, em um primeiro momento, a atenção ou a mediação dessas ferramentas (Lins; Parreiras & Freitas 2021; Segata 2020).

Somada às discussões especificamente relativas à prática etnográfica, testemunhamos, nas duas últimas décadas, uma considerável expansão no número de publicações, bem como de linhas, grupos e centros de pesquisa, e até mesmo departamentos e programas de graduação e pós-graduação voltados à antropologia digital e a áreas correlatas. Uma das abordagens mais recentes aposta justamente na fecundidade de um olhar antropológico dedicado especificamente aos dados (Douglas-Jones; Walford & Seaver 2021), ao que podem fazer ver (MacKenzie & Munster 2019), à sua estética própria (Manovich 2011b; Walford 2021), aos procedimentos exigidos para que se transformem em narrativas significativas (Manovich 2000; Dourish 2004; Dourish & Gómez Cruz 2018), às implicações políticas de seus arranjos técnicos (Ruppert; Isin & Bigo 2017), às infraestruturas que requisitam (Knox 2021; Goldstein & Nost 2022), às filtragens de que dependem (Kockelman 2013).

*Big data*, dados abertos, banco de dados, proteção de dados pessoais, ciência de dados, infraestrutura de dados, justiça de dados. Muita tinta vem sendo gasta e muito dinheiro tem sido investido em pesquisas ao redor de todo o mundo interessadas nesse "novo ecossistema" (boyd & Crawford 2012), em que dados são convertidos em informação (Mirowski & Nik-Khah 2017) a partir de maquinarias que viabilizam o seu processamento adequado – veloz e eficiente – por procedimentos comparativos realizados por máquinas encarregadas da gestão de fragmentos e totalidades de pessoas e coisas em ambientes tecnológicos. Nossa economia e nossos governos estão se tornando orientados por dados, embora ainda estejam em disputa e carentes de efetiva regulação dos seus métodos de coleta e de uso. Parte significativa da produção bibliográfica que responde a esse novo cenário em

fervilhante transformação desponta do que, no Brasil, chamaríamos de estudos sociais aplicados. Pesquisas bem financiadas e desenvolvidas dentro dos próprios Gigantes do *Big Tech* (como na Microsoft Research, na Google Research ou em centros de pesquisa direta ou indiretamente relacionados com empresas produtoras e difusoras de algoritmos, por exemplo). O setor privado se abastece vastamente dos estudos sociais voltados ao funcionamento dos algoritmos, aos usos variados dos computadores, aos padrões de comportamento dos usuários de ferramentas de busca, aplicativos e plataformas digitais.<sup>4</sup>

Mas a eles se soma um conjunto robusto de autores que argumentam ser necessário aos cientistas sociais, filósofos, críticos culturais e teóricos das mídias um posicionamento atento aos impactos dos dados e das ferramentas digitais que leve em conta as relações de poder envolvidas nas novas formas como têm sido organizados nossos sistemas de justiça (Hildebrandt & Rouvroy 2013; Jasanoff 2017; Christin 2017; Wang 2020), as práticas educacionais (Williamson 2015; Ratner 2016), os modelos e os sistemas de saúde (Monteiro 2009; Choroszewicz & Alastalo 2021), os programas de assistência social (Allhutter et al. 2020), a segurança pública e o sistema carcerário (Ferguson 2017; Egbert & Leese 2021), os modelos de administração e gestão pública (Broomfield & Reutter 2022), as obras de infraestrutura (Knox 2016, 2021), a governança ambiental (Goldstein & Nost 2022; Bakker & Ritts 2018). Trabalhos que também discutem os efeitos e os riscos da aproximação entre os algoritmos e a democracia (Rouvroy & Berns 2013; Rouvroy 2017; Mozorov 2018; Zuboff 2018; Cesarino 2019, 2020); que se dedicam às ameaças da *datificação* das vidas e dos governos (Pink & Lanzani 2018; Redden 2018; Segata & Rifiotis 2021; Lemos 2021); aos efeitos nocivos dos vieses algorítmicos (Gillespie 2014; Cesarino 2022); às arquiteturas racistas, misóginas, xenófobas e transfóbicas, por exemplo, que essas ferramentas podem assumir (Ali 2014; Noble 2021; Silva 2022; Faustino & Lippoldi 2023); aos acessos diferenciais às tecnologias e às ferramentas digitais imbricadas nas possibilidades de exercício contemporâneo da cidadania (Pavesi 2017; Parreiras & Macedo 2020; Nemer 2019, 2022); além da recente onda de pesquisas dedicadas às inteligências artificiais.

Em publicações recentes, como o *The Routledge Social Science Handbook of AI* (Elliott 2021), pesquisadores avaliam que, mesmo quando essas tecnologias não alteram paradigmas da produção técnica/tecnológica em si, elas têm gerado consequências profundas, sobretudo de ordem social, como, por exemplo, nos modos como se organizam as políticas governamentais, as regulações legislativas, as relações trabalhistas e o próprio fazer acadêmico e científico (Elliott 2021). Mas, além disso, este cenário de profusão de inteligências artificiais tem promovido discussões importantes a respeito das fronteiras cada vez mais incertas que delimitam a humanidade e a agência de não humanos. No caso específico das máquinas, as análises sugerem que um ponto de partida viável seria traçar como a

agência maquínica tem sido gerada via novas tecnologias digitais e informacionais (Suchman 1998) e como, nesse processo, criam-se tensões na delimitação do que é natural e próprio do humano e o que é artificial, próprio das máquinas. Do fato de essas fronteiras serem pressionadas e indefinidas derivam novos problemas: se não temos claramente estabelecidos os limites entre a inteligência natural e a artificial, de que maneira poderemos atribuir ou compartilhar a responsabilidade acerca das consequências e dos efeitos desse trabalho conjunto e muitas vezes indistinto? Autores como Müller (2022), Crawford (2021) e Coeckelbergh (2024), por exemplo, discutem não apenas as promessas, mas também os riscos envolvidos na autonomia crescente dos sistemas de IA e novas tecnologias de dados informacionais, assim como as questões éticas envolvidas nas formas de controle desses sistemas, que têm passado da condição de objeto à de sujeitos autônomos de interação social.

O que temos visto, portanto, é que os últimos trinta anos consolidaram um campo de estudos sociais e etnográficos em torno da ciência e da tecnologia dos dados digitais que tem ampliado e incitado a antropologia a repensar a divisão fundante entre natureza e cultura, bem como a atualizar possibilidades descritivas e analíticas das relações entre humanidade e entidades outras. Os dados e suas abundantes conexões, que invadem mesmo as etnografias em um primeiro momento desinteressadas das máquinas, de sistemas, aplicativos, informações binárias, imagens georreferenciadas, botões e dispositivos eletrônicos dos mais variados (Douglas-Jones; Walford & Seaver 2021), recolocam de modo muito evidente o tema da comparação na antropologia e as formas como lidamos com nossos próprios dados de pesquisa.<sup>5</sup>

Ao aproximarmos o debate sobre o método comparativo na teoria antropológica das discussões sobre os dados e seu potencial de alterar o fazer etnográfico, é importante nos lembrarmos que, embora os dados digitais possam se assemelhar na aparência, dados semelhantes não se originam necessariamente de lugares semelhantes — e esta tem sido uma preocupação da antropologia muito anterior à Era Digital (Boas 1987). São metodologicamente centrais a qualquer trabalho antropológico reflexões cuidadosas a respeito de quem coletou ou criou os dados; com quais finalidades; a pedido de quem; orientados por quais perguntas e interesses de pesquisa; utilizando que instrumentos. Afinal, como lembra Lisa Gitelman (2013), retomando a ideia de Geoffrey C. Bowker, *raw data* (o dado bruto ou "cru") é um oxímoro, os dados já estão sempre em grande medida "cozidos" ou tratados. Ainda assim, também é necessário enfatizar que, em decorrência de suas importantes singularidades técnicas, há análises que podem ser feitas com dados digitais que sequer haviam sido imaginadas com dados analógicos: mapas dinâmicos, colaboração em tempo real com colegas de todo o mundo e usos interativos de dados — os quais, inclusive, podem se modificar durante o trabalho de escrita e consolidação dos resultados das pesquisas. São novas

*potencialidades* abertas pelos dados digitais que também transformam os modos como conduzimos nossos próprios empreendimentos comparativos e reclamam reflexões cuidadosas a respeito dos alcances do método etnográfico (Beaulieu 2017; Fortun et al. 2017; Albris et al. 2021; Pedersen 2023).

Dessas preocupações derivam outras de igual magnitude e importância, a exemplo das relacionadas ao chamado *colonialismo de dados*, prática que combina o extrativismo predatório do colonialismo histórico com os métodos da quantificação abstrata da computação (Milan & Treré 2019; Cassino et al. 2021; Faustino & Lippold 2023). Entender os *big data* do *Sul Global* significa considerar a atual dependência do capitalismo em relação a esse novo tipo de apropriação que funciona em todos os pontos do espaço onde pessoas ou coisas estão ligadas às infraestruturas de conexão eletrônicas. Assim como o colonialismo histórico, a longo prazo, forneceu as pré-condições essenciais para o desenvolvimento do capitalismo industrial, deparamo-nos com um novo modelo de colonialismo que fornece as pré-condições para um estágio do capitalismo em que a apropriação de todos os aspectos da vida humana sob a lógica e a linguagem dos dados tem sido fundamental (Couldry & Mejias 2018). A etnografia e a reflexão antropológica nos oferecem as ferramentas para compreendermos e resistirmos a essa nova forma de exercício do poder em andamento contemporaneamente (Dencik et al. 2022).<sup>6</sup>

A natureza da pesquisa antropológica (e as implicações derivadas de seu envolvimento direto com as pessoas — no passado, presente e futuro) exigiu que a ética fosse um tema central em toda a discussão sobre os dados, com foco especial no tratamento dos materiais que representam ou afetam aqueles que participam de nossas pesquisas. Nossos textos, áudios, fotografias, filmes e vídeos levam as experiências práticas com dados digitais e incluem discussões sobre precauções éticas a serem consideradas ao apresentarem informações relativas a vidas individuais. Tradicionalmente, a pesquisa em ciências sociais consiste em uma síntese de trabalho de campo e produção de dados — que se influenciam mutuamente — e a situação não é diferente hoje. Como ressalta Pedersen (2023), os emaranhados entre a antropologia e a ciência de dados não podem ser reduzidos a uma relação que coloca, de um lado, o sujeito ético que estuda e, de outro, um objeto êmico de estudo. À medida que os instrumentos de nosso trabalho se tornam mais habilitados, esperados e compartilhados digitalmente, devemos considerar a natureza de nossos dados e seus possíveis usos em um futuro próximo e distante. Neste sentido, Pink e Lanzani (2018) argumentam que uma ética “focada no futuro” como orientação para a prática etnográfica — baseada em seu envolvimento com a incerteza do que ainda é desconhecido, mais do que com uma ética com referência ao passado — pode ajudar a construir análises mais responsáveis do que os cenários preditivos que muitas vezes resultam das análises dos *big data*. Uma ética de pesquisa, assim, que se envolva com os contextos emergentes nos quais

os dados são produzidos e com suas temporalidades; que não seja simplesmente aplicada, mas que faça parte dos processos de pesquisa e análise.

Como procuramos demonstrar nesta breve síntese de uma vasta bibliografia em crescente profusão, este dossiê se insere em meio a debates bastante atuais, politicamente incontornáveis e já bem instalados institucionalmente. Para circunscrever nossa proposta e os objetivos desta publicação, será preciso, no entanto, reforçar um aspecto metodológico relevante: embora se sobreponham e se confundam com frequência em muitos enunciados apressados, o digital, os softwares, os sistemas, os algoritmos, as infraestruturas computacionais, as plataformas, as interfaces, as bases de dados, as mídias, os aplicativos, dentre muitos outros substantivos mobilizados em pesquisas interessadas nessa *ecologia contemporânea particular*, para usar um termo bem formulado por Schinkel (2023), não devem ser considerados exatamente sinônimos. Há entre eles especificidades importantes que, justamente, a etnografia deve se encarregar de esmiuçar. Aí repousa a contribuição fundamental de trabalhos etnográficos, como os reunidos nas páginas que se seguem: na atenção inegociável ao modo como as noções corriqueiramente estabilizadas são postas em contato e em funcionamento em diferentes contextos, por diferentes agentes. Na atenção aos processos – sempre impermanentes – de estabilização de sentidos; na descrição cuidadosa das práticas e das metáforas criadoras.

Neste sentido, destacamos uma vez mais que os artigos a seguir foram instados a refletir especificamente a respeito da rentabilidade de uma atenção antropológica aos dados – e, mais particularmente, aos *big data*. Não nos dedicamos com exclusividade às máquinas, aos códigos, aos comandos eletrônicos. Tampouco prestamos atenção apenas nas pessoas humanas que colecionam, organizam e classificam toda sorte de materiais, combinando-os criativamente e atribuindo-lhes distintos valores com a ajuda de aparatos eletrônicos. Ao olhar para os dados, fomos conduzidas *através* de escalas distintas; transitamos entre graus variados de materialidade; descrevemos processos específicos de tradução. Dedicamo-nos a explorar a polissemia dos dados em diferentes contextos etnográficos e os processos técnicos que os tornam legíveis e comunicáveis. Reconhecer sua socialidade (Walford 2013) e compreender os sistemas de dados como de natureza sociotécnica oferecem uma posição conceitual para desafiar o mantra teleológico da Revolução dos Dados (Kitchin 2021). Ou seja, permite-nos contestar sua inevitabilidade e o modo como os sistemas orientados por dados funcionam, imaginando formas alternativas para os seus usos que explorem com a necessária hesitação, para retomar uma vez mais essa importante imagem de Stengers (2007), questões morais e éticas relativas à forma como são produzidos, compartilhados, comercializados e protegidos.

Em um momento em que dados se tornam uma epistemologia universalizante, que explica tudo e quase nada, unimo-nos à proposta de Douglas-Jones, Walford e Seaver (2021) e procuramos ampliar o coro de trabalhos etnográficos que optam por explorar suas multiplicidades e particularidades realocando-os em diversos mundos locais. Exploramos, com maiores ou menores ênfases, os imaginários, as infraestruturas, as interfaces, as identidades, os interesses e os investimentos (Schinkel 2023) relacionados à ecologia dos dados. Procuramos examinar, sem confundir, sua relação com as novas tecnologias digitais, atenta(o)s aos processos de ruptura, mas também às continuidades das formas analógicas de manejo de informações nos processos de criação, comparação e governo (Douglas-Jones; Walford & Seaver 2021:10-11). Voltamo-nos, enfim, às variadas formas de visibilidade que moldam e são moldadas por dados, bem como às arenas de cálculo que delas decorrem (Douglas-Jones; Walford & Seaver 2021:13). Na seção a seguir, veremos o que propõem os artigos e os diálogos que estabelecem entre si.

### 3

"Se os efeitos coletivos do encantamento e da aceleração atraem a nossa atenção para a inovação, então de quais outros processos eles potencialmente nos desviam?". A pergunta de Katherine Kenny e colegas (2021), em um de seus trabalhos sobre medicina de precisão, é retomada por Evandro Smarieri e Pedro Ferreira ao examinarem um sistema de telemetria utilizado na agricultura brasileira contemporaneamente. Sua indagação se levanta como um ponto de partida apropriado para aproximar e colocar em diálogo os artigos que compõem o dossiê. Para onde tem se voltado nossa atenção quando tratamos de *big data*? Que aspectos as têm arrastado e seduzido, seja como usuários das plataformas, dos bancos de dados, das ferramentas digitais de pesquisa, mas também como antropóloga(o) s direta ou indiretamente, deliberada ou acidentalmente interessada(o)s em compreender seus funcionamentos, seus efeitos e seus alcances culturais, políticos e sociais? Como tem se conformado a economia política que organiza os modos como percebemos e experimentamos o mundo por meio dos dados? (Pedersen; Albris & Seaver 2021).

Talvez já não tão encantados com as seduções da inovação intérmina e ilimitada, talvez receosos e resistentes aos perigos da aceleração, é plausível que nossas análises acabem reduzidas aos aspectos destrutivos da ordem neoliberal capitalista (Deleuze & Guattari 1995-1996; Mirowski 2013). Mas, como afirmam Magda Ribeiro e Manuel Luz a partir da biotecnologia e da economia, a ênfase no aspecto destrutivo tem se mostrado insuficiente como elemento analítico e também, argumentamos, como ferramenta política na escrita etnográfica. Inovação

e aniquilação, sustentam Ribeiro e Luz, andam de mãos dadas e se retroalimentam em procedimentos alternativos de variação e uniformização, de estabilidade e inconstância, de criação e submissão da multiplicidade, em que o gerenciamento de coleções de dados sobre organismos e seres se torna imprescindível. Os autores nos fazem ver como a engenharia genética necessita de repositórios de dados e informações variadas para que possa inovar, ainda que esta inovação tenha como finalidade o extermínio da diversidade, uma vez que se orienta pelo padrão de aniquilação típico do modo de acumulação neoliberal. Este mesmo padrão, argumentam, também tem conduzido as seleções e as transformações do *mainstream* econômico sobre as abordagens heterodoxas da disciplina, caminhando em relação direta com a valorização do capital financeiro e, ao mesmo tempo, com a asfixia do pensamento divergente.<sup>7</sup>

Quando organizamos o dossiê, nosso interesse era o de congregar reflexões a respeito do que os dados, a partir de suas vinculações com aparatos diversos – aplicativos, satélites, microscópios, documentos, algoritmos – são capazes de fazer ver e o que, pelo campo de visualização que criam, autorizam a dizer e a realizar. Eles fazem ver a *terra*, como descreve Vanessa Perin, e são capazes, pelas imagens que geram e pelos modos como elas se organizam, de criar investimentos potenciais. Com Marko Monteiro (2015), a autora afirma que essas ferramentas de visualização dos territórios criam mais do que formas de enxergá-los. Criam imagens-relações, que ajudam a construir esses lugares de forma muito particular. Espelhos, por outro lado, são centrais ao trabalho de Sara Munhoz – a justiça que só se efetiva por meio de Espelhos digitalmente produzidos e postos em circulação –, mas também são evocados por Smarieri e Ferreira, que descrevem como as imagens digitalmente produzidas a partir da coleta e do tratamento de dados não se limitam a representar, mas esculpem a realidade física de maneira dramática: um mundo ordenado para que se aproxime cada vez mais dos pixels, esses pontos fosforescentes que, supostamente, deveriam representá-lo (Hardwood 2008). Vemos tanto na gestão da vida como na dos desejos – com Ribeiro e Luz, mas também com Carolina Parreiras – que as realidades se dão nos interstícios e nos acoplamentos de sistemas homens-máquinas, como argumenta Munhoz, em diálogo com Maurizio Lazzarato (2014). Informações já não exclusivamente antropomórficas, mas elaboradas pela lógica das entradas e saídas, dos *inputs* e *outputs*, encarregados da construção abstrata ou representacional de equivalências. Assim, a *terra*, retomando Perin, imagem-relação decorrente dos aparatos técnicos de manejo dos dados do projeto Embrapa-Moçambique, não é mais apenas lugar de identidade, de pertencimento, de história e de luta. Sua polissemia se expande e se esparrama. Ela é também recurso, é solo, é meio ambiente, é território e paisagem. É, ainda, e isso é absolutamente fundamental, ativo investível, zona

de aptidão, *commodity*, espaço vazio com uma vocação agrícola a ser explorada. Mas, em todos os casos, é *terra* que deve ser lida pelas lentes das *similaridades*.

Como veremos, o artigo de Tone Walford preocupa-se com a concomitância apenas aparentemente paradoxal entre a abertura dos dados e a sua mercantilização. Também aborda o problema do vínculo de sangue entre inovação e aniquilação da vida, das gentes, das diversidades, tal como Ribeiro e Luz o fazem. Enfrenta o problema do "jogo delirante" do capitalismo que, nas palavras de Jorge Villela (2020, 2024), confisca toda vida e toda matéria disponíveis (e há sempre mais meios de tornar tudo, a todo momento, disponível ao confisco) para que sejam convertidas em cifras comercializáveis. Dados, arrancados da terra, da floresta, dos ventos, das águas, da chuva, do plantio, até mesmo dos vazios, que são forçados à redução das equivalências, das similitudes, dos padrões, da univocidade. Confiscados à forma mercadoria, que exige, como vários dos artigos também discutem, esforços constantes para a sua manutenção e cuidado, de modo que deixem de ser vistos como gastos para serem encarados como recursos disponíveis — e, aos poucos, mais do que disponíveis, incontornáveis.

Embora não se desvie desse problema fundamental, Walford propõe não parar aí. A questão levantada em seu artigo é a de como não limitar as nossas próprias descrições a espelhos, representações, ainda outras abstrações de justo aquilo que o próprio mercado e o próprio Estado se empenham em tornar visível: as equivalências desmaterializadas, os fluxos despolitizantes. Como calcar, enraizar, descrever a memória criadora e criativa dos dados, com sua teimosia, sua relutância e sua imponência? Como fazer com que os textos proliferem caminhos alternativos, que mantenham sempre mais dimensões à vista? Como as alternativas também podem seguir contando suas histórias sem serem empurradas, também elas, à opacidade e à neutralidade da abstração e das similitudes?

A viabilidade do domínio político e comercial dos Estados-Nação, como os artigos de Munhoz e de Perin procuram demonstrar junto com uma vasta bibliografia já bastante consolidada, decorre da necessária criação de equivalências. O Estado democrático de direito, sua lei e sua ordem implicam a redução de heterogeneidades pulsantes a sistemas unívocos, a grades classificatórias totalizantes, a plurais sempre provisórios, passíveis de simplificação e de reprodução serial. Esses procedimentos, tão antigos quanto os Estados o são, têm se mostrado cada vez mais dependentes da lógica e da organização de bases de dados digitais. No artigo de Munhoz vemos, por exemplo, como a *uniformização* da lei, exigência absolutamente elementar ao reconhecimento do aspecto democrático de um Estado, só pode se dar a partir de procedimentos bastante custosos e frequentemente obscurecidos de filtragens, comparações e sínteses que envolvem aplicativos computacionais e telas de computadores, mas também o "trabalho pensante" de analistas humanos encarregados de simplificar a linguagem. O que os ministros das cortes superiores



dizem dizer a lei decorre do manejo cuidadoso, humano e maquínico de bancos de dados cada vez mais robustos, ordenados de modo a informar, mas também, e isto é fundamental, a predizer, a influenciar, a persuadir todo o sistema de justiça nacional.

Ainda que os fluxos desterritorializantes do mercado escapem por todos os lados do desejo uniformizador do Estado, Walford discute, em um diálogo com a geografia crítica, como a noção de mercadoria também recorre a esta mesma lógica da produção técnico-política de correspondências, de submissão da diversidade à legibilidade, que cria campos específicos de comparabilidade alimentados, facilitados e promovidos pela lógica e a estética dos dados. Um ponto de abstração suficiente; classificação, agrupamento, ordenamento do heterogêneo para que a circulação — de mercadorias, de dinheiro, de informação, de entendimentos, de regulações — se garanta. Discussão semelhante é apresentada nas etnografias de Smarieri e Ferreira, bem como na de Perin, ao descreverem, cada um a seu modo, os procedimentos técnicos, nada intrínsecos, capazes de transformar a terra e as práticas agrícolas em recurso *potencial* ou em mercadoria. Os autores mostram como, em diferentes contextos, os dados auxiliam na condução de um exercício retórico que legitima a busca incessante de governos e entes privados pelo aumento da produtividade na agricultura. No artigo de Smarieri e Ferreira, a retórica da *precisão* possibilitada pelos dados emerge como estratégia de gestão que vai garantir eficiência no uso de recursos, qualidade, rentabilidade e sustentabilidade da produção agropecuária. Já no caso analisado por Perin, é a retórica das *similaridades* estabelecidas pelos dados disponibilizados pela plataforma WebGIS que vai permitir *desenvolver* a agricultura comercial no norte moçambicano e atrair investimentos. Aparatos maquínicos, algorítmicos, de georreferenciamento, ancorados em enormes bancos de dados, capazes de conjurar imaginários mercadológicos de lucratividade e precisão por meio de suas sistematizações específicas. Ribeiro e Luz, por outro lado, destacam como os procedimentos contemporâneos da produção biotecnológica impactam e orientam o mercado agrícola, absolutamente dependente da lógica das bases de dados para a produção de organismos geneticamente modificados e a subsequente utilização dessas melhorias laboratoriais em sistemas de *plantation*, que aniquilam qualquer possibilidade biodiversa.

Sem rejeitar a importância de nos voltarmos aos procedimentos estabilizadores que circunscrevem e tornam efetivas as práticas de Estado e de mercado, Walford aposta em uma abordagem que coloca temporariamente em suspensão a imagem sedutora e englobante da mercadoria para propor um tropo analítico alternativo, inspirado pelo trabalho de campo, mas também pelo desvio deliberado a um novo *corpus* bibliográfico. Defende um retorno ao território como conceito que possibilita, ao mesmo tempo, englobar as abstrações que sustentam o poder de cálculo dos

dados e os lugares, as narrativas e as tecnologias de extração específicos que "aterram" as coleções de dados. Um tropo, portanto, que se atém às dinâmicas de materialidade, sociabilidade, história e espacialidade, que de forma emergente tanto geram os dados ambientais quanto são produzidas por eles. Nessa direção, ganham força e assumem novos sentidos as imagens de *rastros*, de *mapeamentos* e de *cartografias* mobilizadas em vários dos artigos aqui reunidos. No de Perin, por exemplo, a cartografia é fundamental para a delimitação das *potencialidades* da *terra* e de seus usos políticos, na medida em que sua escala permite colocar em relação territórios singulares como o Cerrado brasileiro e a Savana Moçambicana e ressaltar equivalências que vão legitimar projetos técnicos de *cooperação sul-sul* entre estes dois países. Do mesmo modo, Smarieri e Ferreira nos apresentam às novas tecnologias da agricultura de precisão, cujos mapas produzidos quase em tempo real resultam não apenas na colheita eficiente da lavoura, mas na coleta de dados variados sobre meteorologia, topografia, elementos físico-químicos, informações financeiras, dentre outros, que futuramente irão influenciar as ações no campo. Os modos como Parreiras aciona essas mesmas imagens levanta ainda questões metodológicas que merecem destaque. Por um lado, a própria etnografia é apresentada como "experimento de *mapeamento* e cartografia de plataformas digitais" [grifo nosso], que permite acompanhar os *rastros* deixados pela categoria "novinha" no antigo Twitter. Simultaneamente, a autora afirma que "*a big data permite mapear* conexões e relações entre sujeitos, grupos sociais, tecnologias e estruturas da informação" [grifos nossos]. *Mapeamentos*, então, emergem como imagens "de campo", mas também são mobilizados como procedimentos metodológicos do fazer etnográfico.

Ademais, os artigos do dossiê voltam-se para as máquinas produtoras e gerenciadoras de dados, para sua ética e para a relação que estabelecem com as pessoas que as circundam. Examinam os efeitos muito concretos na terra, na natureza, em humanos e não humanos que se tornam objetos dessas intervenções escalares. Com Parreiras, vemos como os *big data* invadem, informam e influenciam os corpos e os desejos. Investigando as relações entre pornografia e ferramentas de busca digitais, a autora demonstra como *tags* e outras formas de indexação digitais criam personalização, predizem, vigiam, discriminam, moralizam, fixam, direcionam o olhar dos usuários. E, ainda, como sob a aparência da gratuidade, as corporações geram dados sobre dados, constroem cuidadosamente outras métricas, que sempre se beneficiam da opacidade garantida pelas infraestruturas dessas plataformas, e assim se tornam geradoras de riqueza e ordenadoras dos desejos. Métricas que, por sua vez, quando transitam do mercado ao Estado, e vice-versa, tornam-se matéria-prima de equivalências simplificadoras com efeitos políticos muito concretos e perigosos, como a que associa, direta e irreversivelmente, a

categoria "novinha", discutida pela autora, aos discursos moralizantes de um Estado ultraconservador.

A discussão é fundamental e se aproxima daquela proposta por trabalhos à primeira vista muito distantes, como o de Munhoz, que não trata de plataformas digitais privadas nem aborda a produção e a circulação de desejos sexuais, mas também se depara com a força descomunal das formas de classificação e indexação em ferramentas de busca digitais, bem como com o desejo onipresente e inalcançável da transparência. Ambos os artigos, ao examinarem documentos técnicos (simultaneamente acessíveis e inapreensíveis) e ao descreverem procedimentos de organização das bases de dados em uma interface envolvente e intuitiva, enfrentam etnograficamente os arranjos ópticos específicos que ocultam, peneiram, selecionam, orientam e fixam determinados padrões sob abundâncias tecnicamente domesticadas. Smarieri e Ferreira, por sua vez, discutem como a exigência da precisão na agricultura transforma toda ação em um rastro digital passível de processamento, que, ao mesmo tempo, esbarra em gargalos infraestruturais e humanos. "É natural ser digital", propagandeia o sistema de telemetria. Mas para que as máquinas informem e orientem adequadamente, é preciso investir. Inclusive nas pessoas, que não são naturalmente digitais. O que nos faz retornar ao trabalho de Parreiras que, acompanhando os argumentos de Hine (2000) e de Seaver (2022), enfatiza a importância de recolocar o sujeito em meio aos conjuntos de dados, percebendo os fluxos de formas de classificação e significados, associação de categorias e sentidos distintos. Como alerta a autora, uma antropologia dedicada aos dados, ainda que atenta aos seus maquinismos, não pode perder de vista as pessoas, e a maneira como influenciam e moldam o funcionamento dos artefatos.

Os artigos aqui reunidos compartilham a especificidade de não descreverem, exatamente, as práticas dos usuários das plataformas, os modos como as pessoas se apropriam das infraestruturas digitais que lhes dão acesso (sempre mediado, significativamente moldado e limitado) aos dados. Em geral, as(os) autoras(es) dos artigos fincam-se em outra escala que também permite "povoar" os ecossistemas de dados: descrevem os projetos institucionais que dão origem aos aparatos infraestruturais que os classificam e organizam; acompanham os desejos das empresas e dos governos de criarem novos modos de ver, comparar e governar. Apresentam, é verdade, os trabalhos técnicos e as tarefas de alguns profissionais encarregados do manejo direto desses aparelhos cuja interface lhes garante a relação com a prometida eficiência do mundo dos dados. Uma atenção aos trabalhos ocultos sob as atrativas interfaces das plataformas, ferramentas de busca, imagens digitais das mais variadas: sem as infraestruturas, nada pode ser comparado ou criado. O acúmulo de dados não garante sua legibilidade e acessibilidade, é preciso certificar tecnicamente a conectividade, na mesma medida em que é

preciso assegurar que dados "brutos" (e ainda que nunca totalmente "brutos") sejam tornados significativos.

O que, afinal, os dados criam ao embaralharem humanos e não humanos, dígitos e mapas, sensores e políticas públicas, desejos e repressões, ordem e resistências, inovação e aniquilação, acesso e opacidade, democracia e mercado? Ao falarmos de justiça, democracia, neoliberalismo, cooperações transnacionais, plataformas digitais milionárias; ao mobilizarmos mapas, técnicos, imagens pixeladas, pastas de documentos, algoritmos, telas, tratores, sementes, vídeos com o intuito de trazermos à luz os caminhos por onde os dados circulam, os procedimentos comparativos que ensejam e os desejos de governança que promovem, convidamos os leitores, uma vez mais, a uma atenção cuidadosa e permanente aos efeitos desses complexos "sistemas de extração" (Crawford & Joler 2020), de escala planetária, absolutamente palpáveis, territoriais, materiais, que não podem ser negligenciados sob a aparente inapreensão sedutora dos dados digitais.<sup>8</sup>

Em 2023, o Congresso Nacional brasileiro votou e aprovou, com ampla maioria, a PL 490/07, conhecida como o Projeto do Marco Temporal, que afirma que os povos indígenas só teriam direito à demarcação de suas terras caso já as estivessem ocupando no dia 5 de outubro de 1988, data da promulgação da Constituição brasileira. As discussões foram acaloradas e envolveram também o Supremo Tribunal Federal, que decidiu pela inconstitucionalidade da tese. O debate ainda não está completamente encerrado. É importante não desvincularmos essas discussões sobre os *big data* das lutas territoriais que marcam toda a história colonial brasileira, mas que têm se intensificado, muito concretamente, nos últimos anos. Do que trata essa PL senão de um esforço deliberado de redução da terra à forma-mercadoria, às métricas que tornam visíveis e concretas suas potencialidades mercadológicas enquanto silenciam, sob a linguagem-comando da justiça, toda a volumetria que a enforma? Os artigos a seguir nos oferecem mais um caminho possível para encontrarmos ferramentas de luta, em nossa escrita, mas não apenas nela, que mantenham abertos os caminhos de resistência aos padrões da aniquilação.

## Notas

- 1 Disponível em: <https://itsrio.org/pt/artigos/por-que-sam-altman-foi-demitido-da-open-ai/> e <https://www.cnnbrasil.com.br/economia/negocios/sam-altman-retorna-ao-conselho-de-administracao-da-openai-empresa-por-tras-do-chatgpt/>. Acesso em 02/05/2024.
- 2 A atenção das ciências sociais às máquinas não é recente. Para uma cuidadosa revisão a respeito dessa relação nos XIX e XX, usualmente ocultada da história das disciplinas, cf. Brito (2019).
- 3 Dois exemplos recentes ressoam nesta proposta. Em 2017, Katherine Cook publicou uma série de *posts* em defesa da aproximação de antropólogas e arqueólogas das linguagens de programação, e a incorporação efetiva de aplicações digitais seja em sala de aula, ou no desenvolvimento das pesquisas acadêmicas antropológicas. Disponível em: <http://www.utpteachingculture.com/coding-culture-why-anthropology-students-and-their-instructors-should-learn-to-code/>. Acesso em: 02/05/2024. Guilherme Orlandini Heurich, por sua vez, tem liderado um projeto que denominou "code::anth", em uma tentativa de combinar codificação e antropologia. Interessa-se pelas interseções possíveis do desenvolvimento de *software* com as pesquisas com pessoas. Disponível em: <http://code-anth.xyz/>. Acesso em 02/05/2024. Publicou, recentemente, o livro *Coderspeak: The language of computer programmers* (Heurich 2024), sobre a linguagem de programação Ruby e sua comunidade.
- 4 Tarleton Gillespie e Nick Seaver, da Cornell University e da Tufts University, respectivamente, organizaram uma extensa lista com publicações da sociologia, antropologia, estudos da tecnologia, geografia, comunicação, *media studies*, estudos legais etc. para "catalogar a emergência dos 'algoritmos' como objetos de interesse de disciplinas para além da matemática, da ciência da computação e da engenharia de *softwares*" (tradução nossa). Atualizada pela última vez no final de 2016, a lista reúne mais de 200 trabalhos e está disponível on-line em livre acesso em <https://socialmediacollective.org/reading-lists/critical-algorithm-studies/#1.2>. Acesso em 02/05/2024. De lá para cá, a produção se ampliou significativamente. Os organizadores desta lista são pesquisadores vinculados ao Social Media Collective, uma rede de cientistas sociais e pesquisadores das humanidades fundada em 2009, que faz parte dos laboratórios da Microsoft Research, interessados, por sua vez, em "acelerar descobertas científicas e inovação tecnológica para capacitar todas as pessoas e organizações no planeta para alcançar mais" (tradução nossa). Disponível em: <https://www.microsoft.com/en-us/research/about-microsoft-research/>. Acesso em 02/05/2024.
- 5 Se a comparação passou por um profundo ceticismo no início da disciplina, desde que Franz Boas apontou para as severas limitações do método em resposta à Taylor (Schnegg 2014), ela foi revigorada de diversas maneiras, sobretudo quando aliada ao trabalho etnográfico. No conhecido ciclo de seminários *The History of Cross-Cultural Comparativism*, realizados na Universidade de Cambridge, diversos antropólogos, entre eles Philippe Descola, Marilyn Strathern, Carlo Severi, Caroline Humphrey, se reuniram em torno do projeto que buscou comparar culturas em diferentes períodos e

reconsiderar as especificidades da abordagem comparativa em seus diversos momentos históricos. Conferir: <https://www.crassh.cam.ac.uk/research/projects-centres/the-history-of-cross-cultural-comparatism/>

- 6 A etnografia tem se mostrado um procedimento potente, por exemplo, para projetos interdisciplinares que visam fomentar debates em torno da ideia de *data justice*, expandindo tanto agendas de pesquisa quanto formas de ativismos e *advocacy* relacionadas à conexão cada vez mais profunda entre processos de *datificação* e as possibilidades de justiça social. Sobre estas iniciativas, conferir <https://www.datajusticeincontext.com/> e <https://datajusticelab.org/>
- 7 Zuboff (2018), tem sido incansável em alertar para o impacto que a fusão de empresas de tecnologia exerceu nas questões de privacidade de dados sobre preferências e comportamentos e provoca: "como essa mudança impacta a democracia e as possibilidades de futuro na era digital?". Para a autora, o capitalismo de vigilância reivindica a experiência humana como matéria-prima para a tradução de comportamentos em dados, em uma forma de mercado que não se resume ao digital, mas atravessa e transforma as mais amplas esferas sociais.
- 8 O impressionante trabalho "Anatomy of an AI", de Crawford e Joler (2020), é descrito por Cristina Gonzales e Pedro P. Ferreira como uma cuidadosa "cartografia das infraestruturas do infocapitalismo", imagem que emprestamos neste ponto do artigo. O mapa está disponível em: <https://anatomyof.ai/>. Acesso em 02/05/2024.

## Referências

- ALBRIS, Kristoffer; OTTO, Eva; ASTRUPGAARD, Sofie; GREGERSEN, Emilie; JØRGENSEN, Laura; JØRGENSEN, Olivia; SANDBYE, Clara Rosa & SCHØNNING, Signe. 2021. "A view from anthropology: Should anthropologists fear the data machines?". *Big Data & Society*, July-December: 1-7.
- ALABDULLAH, Bayan; BELOFF, Natalia & WHITE, Martin. 2018. "Rise of Big Data – Issues and Challenges". Conference Paper. 21st Saudi Computer Society National Computer Conference (NCC).
- ALI, Syed Mustafa. 2014. "Towards a Decolonial Computing". In: E. Buchanan; P. de Laat; H. Tavani; J. Klucarich (eds.), *Ambiguous Technologies: Philosophical Issues, Practical Solutions, Human Nature*, Anais da International Society of Ethics and Information Technology, Portugal. p. 28-35.
- ALLHUTTER, Doris; CECH, Florian; FISHER, Fabian; GRILL, Gabriel & MAGER, Astrid. 2020. "Algorithmic Profiling of Job Seekers in Austria: How Austerity Politics Are Made Effective". *Frontiers in Big Data*, 3 (5):1-17.
- BAKKER, Karen & RITTS, Max. 2018. "Smart Earth: A meta-review and implications for environmental governance". *Global Environmental Change*, 52:201-211.
- BATESON, Gregory. 1972. *Steps to an Ecology of Mind. Collected Essays in Anthropology, Psychiatry, Evolution and Epistemology*. Northvale, New Jersey, Londres: Jason Aronson Inc.
- BEAULIEU, Anne. 2017. "Vectors for fieldwork: Computational thinking and New modes of ethnography". In: Larissa Hjoth; Heather Horst; Anne Galloway & Genevieve Bell, *The Routledge Companion to Digital Ethnography*. New York: Routledge. p. 55-65.
- BOAS, Franz. 1987. *Anthropology and Modern Life*. New York: Dover Publications.
- BOELLSTORFF, Tom. 2013. "Making big data, in theory". *First Monday*, 18 (10):1-17.
- BOELLSTORFF, Tom & MAURER, Bill. 2015. *Introduction*. In: Genevieve Bell; Tom BOELLSTORFF; Melissa Gregg; Bill Maurer & Nick Seaver. *Data: Now bigger and better!*. Chicago: Prickly Paradigm Press.
- BOYD, Danah & CRAWFORD, Kate. 2012. "Critical Questions for Big Data". *Information, Communication and Society*, 15 (5):662-679.
- BRITO, Rainer Miranda. 2019. "Sobre uma Ciência Social das Máquinas". *Revista Brasileira de Ciências Sociais*, 34:1-18.
- BROOMFIELD, Heather & REUTTER, Lisa. 2022. "In search of the citizen in the datafication of public administration". *Big Data & Society*, 9 (1):1-14.
- CASSINO, João Francisco; SOUZA, Joyce Souza & SILVEIRA, Sérgio Amadeu (orgs.). 2021. *Colonialismo de dados: como opera a trincheira algorítmica na guerra neoliberal*. São Paulo: Autonomia Literária.
- CESARINO, Leticia. 2019. "On digital populism in Brazil". *Political and Legal Anthropology Review – Ethnographic Explainers*.
- CESARINO, Leticia. 2020. "Como vencer uma eleição sem sair de casa: a ascensão

- do populismo digital no Brasil". *Internet & Sociedade*, 1:92-120.
- CESARINO, Letícia. 2021. "Antropologia digital não é etnografia: explicação cibernética e transdisciplinaridade". *Civitas*, 21 (2):304-315.
- CESARINO, Letícia. 2022. *O mundo do avesso: Verdade e política na era digital*. São Paulo: Ubu Editora.
- CHRISTIN, Angèle. 2017. "Algorithms in practice: comparing web journalism and criminal justice". *Big Data & Society*, 1-14.
- CHOROSZEWICZ, Marta & ALASTALO, Marja. 2021. "Organizational and professional hierarchies in a data management system: public – private collaborative building of public healthcare and social services in Finland". *Information, Communication & Society*.
- COECKELBERGH, Mark. 2024. *Ética na Inteligência Artificial*. São Paulo: Ubu Editora.
- COLEMAN, Gabriella. 2010. "Ethnographic Approaches to Digital Media". *Annual Review of Anthropology*, 39:487-505.
- COULDRY, Nick & MEJIAS, Ulises. 2018. *Data colonialism: rethinking big data's relation to the contemporary subject*. Londres: SAGE Publications.
- CRAWFORD, Kate. 2021. *Atlas of AI: Power, Politics, and the Planetary Costs of Artificial Intelligence*. New Haven: Yale University Press.
- CRAWFORD, Kate & JOLER. 2018. *Anatomy of an AI*. Disponível em: <https://anatomyof.ai/>. Acesso em 14/05/2024.
- DELEUZE, Gilles & GUATTARI, Félix. 1995-1997. *Mil Platôs. Capitalismo e Esquizofrenia*. Rio de Janeiro: Editora 34.
- DENCIK, Lina; HINTZ, Arne; REDDEN, Joanna & TRERÉ, Emiliano. 2022. *Data Justice*. Londres: Sage Publications.
- DIEBOLD, Francis X. 2019. "On the Origin(s) and Development of the Term 'Big Data'". *SSRN PIER Working Paper*, 12-037. Disponível em: <https://ssrn.com/abstract=2152421> or <http://dx.doi.org/10.2139/ssrn.2152421> Acesso em 14/05/2024.
- DOUGLAS-JONES, Rachel; WALFORD, Tone & SEAVER, Nick. 2021. "Introduction: Towards an Anthropology of Data". *JRAI Journal of Royal Anthropological Institute*, NS (27):9-25.
- DOURISH, Paul. 2004. "What we talk when we talk about context". *Personal and Ubiquitous Computing*, 8 (1):19-30.
- DOURISH, Paul & GÓMEZ CRUZ, Edgar. 2018. "Datafication and data fiction: Narrating data and narrating with data". *Big Data and Society*, julho-dezembro:1-10.
- EGBERT, Simon & LEESE, Matthias. 2021. *Criminal Futures. Predictive Policing and Everyday Police Work*. London and New York: Routledge.
- EISENLOHR, Patrick. 2004. "Language revitalization and new technologies: cultures of electronic mediation and the refiguring of communities". *Annual Review of Anthropology*, 33:21-45.
- ELLIOT, Anthony (ed.) 2022. *The Routledge Social Science Handbook of AI*. Oxon e New York: Routledge.
- ESCOBAR, Pablo. 1994. "Welcome to Cyberia. Note on the Anthropology of Cyberculture". *Current Anthropology*, 35 (3):211-231.
- FAUSTINO, Deivison & LIPPOLD, Walter. 2023. *Colonialismo digital. Por uma*



- crítica hacker-fanoniana*. São Paulo: Boitempo.
- FERGUSON, Andrew Guthrie. 2017. *The Rise of Big Data Policing: Surveillance, Race, and the Future of Law Enforcement*. New York: New York University Press.
- FERREIRA, Pedro P. 2006. "Transe maquínico ou: o que pode uma máquina?". *NADA*, Lisboa, 8:74-77.
- FORTUN, Mike; FORTUN, Kim & MARCUS, George. 2017. *Computers in/and anthropology: The Poetics and Politics of Digitization*. In: Larissa Hjoth; Heather Horst; Anne Galloway & Genevieve Bell, *The Routledge Companion to Digital Ethnography*. New York: Routledge. p. 11-20.
- FULLER, Matthew. 2006. *Behind the Blip: essays on the culture of software*. Brooklyn, NY: Autonomedia.
- FULLER, Matthew. 2008. *Software Studies: a lexicon*. Cambridge, MA: MIT Press.
- GALLOWAY, Alexander. 2004. *Protocol. How Control Exists after Descentralization*. Cambridge, London: The MIT Press.
- GANDOMI, Amir & HAIDER, Murtaza. 2015. "Beyond the hype: Big data concepts, methods, and analytics". *International Journal of Information Management*, 35:137-144.
- GEISMAR, Haidy & KNOX, Hanna. 2012. *Digital Anthropology*. New York: Routledge.
- GILLESPIE, Tarleston. 2014. "The Relevance of Algorithms". In: T. Gillespie; P. J. Boczkowski & K. A. Foot. *Media technologies: essays on communication, materiality, and society*. Cambridge, MA: MIT Press.
- GITELMAN, Lisa (ed.). 2013. *"Raw data" is an oxymoron*. Massachusetts: The MIT Press.
- GOLDSTEIN, Jenny & NOST, Eric. 2022. "Introduction: Infrastructuring Environmental Data". In: \_\_\_\_\_, (eds.). *The Nature of Data: Infrastructures, Environments, Politics*. Lincoln: University of Nebraska Press.
- HARWOOD, Graham. 2008. "Pixel". In: Matthew Fuller (ed.), *Software studies – a lexicon*. Cambridge: MIT Press. pp. 213-7.
- HAYLES, N. Katherine. 1999. *How We Became Posthuman. Virtual Bodies in Cybernetics, Literature, and Informatics*. Chicago: University of Chicago Press Journals.
- HILDEBRANDT, Mireille & ROUVROY, Antoinette (eds.). 2013. *Law, Human Agency, and Autonomic Computing: the philosophy of Law meets the philosophy of technology*. Abingdon: Routledge.
- HYMES, Dell H. 1965. *The Use of Computers in Anthropology*. Mouton: The Hague.
- HINE, Christine. 2005. *Virtual Methods: issues in social research on the internet*. Oxford and New York: Berg Publications.
- HINE, Christine. 2015. *Ethnography for the Internet. Embedded, embodied and everyday*. London/New York: Bloomsbury.
- HORST, Heather A. & MILLER, Daniel. 2012. *Digital Anthropology*. New York: Berg Publications.
- JASANOFF, Sheila. 2017. "Virtual, visible, and actionable: Data assemblages and the sightlines of justice". *Big Data & Society*, 4 (2).

- KENNY, Katherine; BROOM, Alex; PAGE; Alexander; PRAINSACK, Barbara; WAKEFIELD, Claire E.; LWIN, Zarnie & KHASRAW, Mustafa. 2021. "A sociology of precision-in-practice: the affective and temporal complexities of everyday clinical care". *Sociology of Health & Illness*, 43:2178-95.
- KITCHIN, Rob. 2021. *Data Lives. How Data are made and shape our World*. Bristol: Bristol University Press.
- KITCHIN, Rob & MCARDLE, Gavin. 2016. "What makes Big Data, Big Data? Exploring the ontological characteristics of 26 datasets". *Big Data & Society*, 1-10.
- KITTLER, Friedrich. 2017. "Software não existe". In: F. Kittler, *A verdade do mundo técnico: Ensaios sobre a genealogia da atualidade*. Rio de Janeiro: Contraponto.
- KOCKELMAN, Paul. 2013. "The anthropology of an equation". *HAU: Journal of Ethnographic Theory*, 3 (3):33-61.
- KNOX, Hannah. 2016. "An Infrastructural Approach to Digital Ethnography: Lessons from the Manchester Infrastructures of Social Change Project". In: L. Hjorth; H. Horst; A. Galloway & G. Bell, *The Routledge Companion to Digital Ethnography*. New York: Routledge.
- KNOX, Hannah. 2021. "Traversing the Infrastructures of Digital Life". In: H. Geismar & H. Knox (eds.), *Digital Anthropology*. New York: Routledge.
- KNOX, Hannah & NAFUS, Dawn (eds.). 2018. *Ethnography for a data-saturated world*. Manchester: Manchester University Press.
- KOZINETS, Robert. 2010. *Netnography: doing ethnographic research online*. Londres: Sage Publications.
- LANEY, Doug. 2001. "3D data management: Controlling data volume, velocity and variety. Meta Group". Disponível em <http://blogs.gartner.com/douglaney/files/2012/01/ad949-3D-Data-Management-Controlling-Data-Volume-Velocity-and-Variety.pdf>. Acesso em 13/05/2024.
- LAZZARATO, Maurizio. 2014. *Signos, Máquinas, Subjetividades*. São Paulo: Edições SESC / n-1 edições.
- LEMOS, André. 2021. "Dataficação da vida". *Civitas*, 21 (2):193-202.
- LÉVY, Pierre. 2003. *A inteligência coletiva: por uma antropologia do ciberespaço*. São Paulo: Loyola.
- LÉVY, Pierre. 2010. *Cibercultura*. São Paulo: Editora 34.
- LINS, Beatriz Accioly; PARREIRAS, Carolina & FREITAS, Eliane Tânia. 2020. "Introdução ao Dossiê Estratégias para pensar o digital". *Cadernos de Campo*, 29 (2).
- LO AND BEHOLD: *Reveries of the Connected World*. 2016. Direção: Werner Herzog, Estados Unidos.
- LOHR, Steve. 2013. "The origins of 'Big Data': An etymological detective story". *The New York Times*. Disponível em: <http://bits.blogs.nytimes.com/2013/02/01/the-origins-of-big-data-an-etymological-detective-story/>. Acesso em 13/05/2024.
- LUPTON, Deborah. 2015 "The thirteen Ps of big data". *The Sociological Life*. Disponível em: <https://simplysociology.wordpress.com/2015/05/11/the-thirteen-ps-of-big-data/> Acesso em 15/05/2024.
- MACKENZIE, Adrian & MUNSTER, Anna. 2019. "Platform Seeing: image ensembles

- and their invisibilities". *Theory, Culture & Society*, 0 (0):1-20.
- MALABY, Thomas. 2009. *Making Virtual Worlds: Linden Lab and Second Life*. Ithaca: Cornell University Press.
- MANOVICH, Lev. 2000. "Database as a genre of New Media". *AI&Society*, 14:176-183.
- MANOVICH, Lev. 2002. *The Language of New Media*. Cambridge, MA: MIT Press.
- MANOVICH, Lev. 2003. "New Media from Borges to HTML". In: N. Wardrip-Fruin & N. Montfort. *The New Media Reader*. Cambridge, MA: The MIT Press.
- MANOVICH, Lev. 2011a. "Trending: the promises and the challenges of big social data". *Debates in the Digital Humanities*. Disponível em: [http://www.manovich.net/DOCS/Manovich\\_trending\\_paper.pdf](http://www.manovich.net/DOCS/Manovich_trending_paper.pdf). Acesso em 13/05/2024.
- MANOVICH, Lev. 2011b. "What is Visualization?". *Visual Studies*, 26 (1):36-49.
- MANOVICH, Lev. 2013. *Software takes Command*. New York: Bloomsbury Academic.
- MILAN, Stefania & TRERÉ, Emiliano. 2019. "Big Data from the South(s): Beyond Data Universalism". *Television & New Media*, 20 (4):319-335.
- MILLER, Daniel. 2001. *The Internet. An Ethnographic Approach*. London and New York: Routledge.
- MILLER, Daniel & SLATER, Don. 2000. *The internet: the ethnographic approach*. London: Routledge.
- MIROWSKI, Philip. 2013. *Never Let a Serious Crisis Go to Waste*. Londres: Verso Books.
- MIROWSKI, Philip & NIK-KHAH, Edward. 2017. *The Knowledge We Have Lost in Information: The History of Information in Modern Economics*. Oxford: Oxford University Press.
- MONTEIRO, Marko Synésio Alves. 2009. "Digitalizando o câncer de próstata: pensando as interseções entre engenharia e biologia na ciência contemporânea". *Revista de Antropologia*, 52:247-288.
- MONTEIRO, Marko Synésio Alves. 2015. "Construindo imagens e territórios: pensando a visualidade e a materialidade do sensoriamento remoto". *História, Ciências, Saúde – Manguinhos*, Rio de Janeiro, 22 (2):577-591.
- MOZOROV, Evgeny. 2018. *Big Tech: a ascensão dos dados e a morte da política*. São Paulo: Ubu Editora.
- MÜLLER, Vicent. 2022. *Ethics of Artificial Intelligence*. In: Anthony Elliot (ed.), *The Routledge Social Science Handbook of AI*. Oxon e New York: Routledge. p. 122-137.
- NEMER, David. 2019. "Repensando as Desigualdades Digitais: as promessas da Web 2.0 para os Marginalizados". *Revista Tecnologia e Sociedade*, 15 (35).
- NEMER, David. 2022. *Technology of the Oppressed: Inequity and the digital mundane in favelas of Brazil*. Cambridge, MA: MIT Press.
- NOBLE, Safiya Umoja. 2021. *Algoritmos da Opressão. Como o Google fomenta e lucra com o racismo*. Santo André, SP: Rua do Sabão.
- PAVESI, Patrícia. 2017. "A gambiarra, o acesso à internet e a ciência de várzea: consumo de Tecnologias de Informação e epistemologias populares". *Sinais*, 21:323-341.

- PARREIRAS, Carolina. 2008. *Sexualidade no ponto.com: espaços e homossexualidades a partir de uma comunidade on-line*. Dissertação de Mestrado em Antropologia Social, Instituto de Filosofia e Ciências Humanas, Universidade de Campinas – Unicamp.
- PARREIRAS, Carolina & MACEDO, Renata Mourão. 2020. "Digital Inequalities and Education in Brazil during the COVID-19 Pandemic: a brief reflection on the challenges of remote learning". *Digital Culture & Education*, 1:1.
- PEDERSEN, Morten Axel. 2023. "Towards a Machinic Anthropology". *Big Data & Society*, January-June:1-9.
- PEDERSEN, Morten Axel; ALBRIS, Kristoffer & SEAVER, Nick. 2021. "The Political Economy of Attention". *Annual Review of Anthropology*, 50:309-325.
- PINK, Sarah; HORST, Heather; POSTILL, John; HJORTH, Larissa; LEWIS, Tania & TACCHI, Jo. 2016. *Digital Ethnography: Principles and Practices*. London: Sage Publications.
- PINK, Sarah & LANZENI, Debora. 2018. "Future Anthropology Ethics and Datafication: Temporality and Responsibility in Research". *Social Media + Society*: 1-9.
- RABINOW, Paul. 1996. *Making PCR a story of biotechnology*. Chicago: University of Chicago.
- RAMOS, Jair de Souza & FREITAS, Eliane Tânia. 2017. "Etnografia digital". *Revista Antropológica*, 42:8-15.
- RATNER, Helene. 2016. "Databasing Danish Schools". *Theorizing the Contemporary, Fieldsights*, March 24. Disponível em: <https://culanth.org/fieldsights/databasing-danish-schools>. Acesso em 05/07/2024.
- REDDEN, Joanna. 2018. "Democratic governance in an age of datafication: lessons from mapping government discourses and practices". *Big Data and Society*.
- RIFIOTIS, Theophilos. 2016. *Etnografia no ciberespaço como "repovoamento" e explicação*. In: Jean Segata & Theophilos Rifiotis, *Políticas etnográficas no campo da cibercultura*. Brasília: ABA Publicações. pp. 129-152.
- RIFIOTIS, Theophilos; MÁXIMO, Maria Elisa; LACERDA, Juciano & SEGATA, Jean (orgs.). 2010. *Antropologia do ciberespaço*. Florianópolis: Editora UFSC.
- ROUVROY, Antoinette. 2017. "Gouverner hors les norms: la gouvernementalité algorithmique". *Lacan Quotidien*, 733.
- ROUVROY, Antoinette & BERNS, Thomas. 2013. "Gouvernementalité algorithmique et perspectives d'émancipation. Le disparate comme condition d'individuation par la relation?". *Réseaux*, 177 (1):163-196.
- RUPPERT, Evelyn; ISIN, Engin & BIGO, Didier. 2017. "Data politics". *Big Data and Society*, julho-dezembro:1-7.
- RUYER, Raymond. 1954. *La cybernétique et l'origine de l'information*. Paris: Éditeur Flammarion.
- SCHINKEL, Willem. 2023. "Steps to an Ecology of Algorithms". *Annual Review of Anthropology*, 52:171-186
- SCHNEGG, Michael. 2014. "Anthropology and Comparison". *Zeitschrift für Ethnologie*, 139 (2):55-72.

- SEEVER, Nick. 2014. "Structuralism: Thinking with Computers". *Savage Minds*. Disponível em: <https://savageminds.org/2014/05/21/structuralism-thinking-with-computers/>. Acesso em 04/06/2024.
- SEEVER, Nick. 2018. "What should an Anthropology of Algorithms do?". *Cultural Anthropology*, 33 (3):375-385.
- SEEVER, Nick. 2022. *Computing Taste. Algorithms and the Makers of Music Recommendation*. Chicago: University of Chicago Press.
- SEGATA, Jean. 2020. "A colonização digital do isolamento". *Cadernos de Campo*, 29 (1):163-171.
- SEGATA, Jean; RIFIOTIS, Theophilos. 2016. *Políticas etnográficas no campo da cibercultura*. Brasília: ABA Publicações.
- SEGATA, Jean; RIFIOTIS, Theophilos. 2016. *Políticas etnográficas no campo da cibercultura*. Brasília: ABA Publicações.
2021. "Digitalização e dataficação da vida". *Civitas*, 21 (2):186-192.
- SEGATA, Jean; MÁXIMO, Maria Elisa & BALDESSAR, Maria José. 2012. *Olhares sobre a cibercultura*. Florianópolis: CCE/UFSC
- SILVA, Tarcízio. 2022. *Racismo algorítmico: inteligência artificial e discriminação nas redes digitais*. São Paulo: Edições SESC.
- SUCHMAN, Lucy. 1998. "Human/Machine Reconsidered". *Cognitive Studies: Bulletin of the Japanese Cognitive Science Society*, 5 (1):5-13.
- STENGERS, Isabelle. 2007. "La proposition cosmopolitique". In: J. Lolive & O. Soubeyran (eds.), *L'émergence des cosmopolitiques*. Paris: La Découverte.
- STRATHERN, Marilyn. 2017. "A relação: acerca da complexidade e da escala". In: \_\_\_\_\_. *O Efeito Etnográfico e outros ensaios*. São Paulo: UBU. pp. 263-295.
- STRATHERN, Marilyn. 2021. "Counting Generation(s)". *Interdisciplinary Science Reviews*, 46 (3):286-303.
- TSING, Anna. 2019. *Viver nas Ruínas: Paisagens Multiespécie no Antropoceno*. Brasília: IEB MIL Folhas.
- VALENTIM, Julio & PAVESI, Patrícia P. 2021. "Ciências Sociais Computacionais: um novo paradigma para as Ciências Sociais?". *SIMBIÓTICA*, 8:1-16.
- VILLELA, Jorge. 2020. "Confiscações, Lutas Anti-Confiscatórias e Antropologia Modal". In: J. Villela & S. de A. Vieira (orgs.), *Insurgências, ecologias dissidentes e antropologia modal*. Goiânia: Editora da Imprensa Universitária.
- VILLELA, Jorge. "Solenidade e Vida. Por uma composição que nos seja favorável". In: Hippolyte Brice Sogbossi & Leonardo Leal Esteves (orgs.), *Entre distopias e decolonialidades: antropologia em cenários cambiantes*. Aracaju: Criação Editora. pp. 15-37.
- VIRILIO, Paul. 1977. *Vitesse et politique*. Paris: Galilée.
- VISMANN, Cornelia & KRAJEWSKI, Markus. 2007. "Computer Juridisms". *Grey Room*, 29:90-109.
- WALFORD, Tone. 2013. *Transforming Data: An Ethnography of Scientific Data from the Brazilian Amazon*. Tese em Anthropology of Science/STS, University of Copenhagen, Copenhagen.
- WALFORD, Tone. 2017. "Raw Data: Making Relations Matter". *Social Analysis*, 61 (2):65-80.

- WALFORD, Tone. 2018. "If Everything is Information": Archives and Collecting on the Frontiers of Data-Driven Science. In: Hannah Knox & Dawn Nafus (eds.), *Ethnography in a Data-Saturated World*. Manchester: University of Manchester Press.
- WALFORD, Tone. 2021. "Data - Ova - Gene - Data". *Journal of the Royal Anthropological Institute*, 27 (S1):127-141.
- WANG, Nu. 2020. "'Black Box Justice'. Robot Judges and AI-based Judgment Processes in China's Court System". In: *2020 IEEE International Symposium on Technology and Society (ISTAS)*:58-65.
- WIENER, Norbert. 2017. *Cibernética ou controle e comunicação no animal e na máquina*. São Paulo: Perspectiva.
- WILLIAMSON, Ben. 2015. "Governing Software: networks, databases and algorithmic power in the digital governance of public education". *Learning, Media and Technology*, 40 (1):83-105.
- ZUBOFF, S. 2018. *A Era do Capitalismo de Vigilância: a luta por um futuro humano na nova fronteira do poder*. São Paulo: Intrínseca.

**BIG DATA: MODOS DE FAZER,  
COMPARAR E GOVERNAR****Resumo**

Neste dossiê, congregamos pesquisas que descrevem e analisam relações constituídas sob o efeito dos *big data*. A partir de espaços bastante distintos entre si — o jurídico, o burocrático, o agrícola, o biotecnológico, o econômico, o pornográfico —, voltamo-nos às agências conflitivas e criativas implicadas no que os *dados* fazem ver e no que são capazes de criar por meio de seu potencial de hiper-relacionalidade. Discutimos, então, a relação dos volumosos bancos de dados com aparatos diversos, como aplicativos, satélites, microscópios, documentos e algoritmos, capazes de arranjar e distribuir elementos informacionais por meio de práticas comparativas. Examinamos procedimentos de coleta e processamento de *big data* em diferentes setores e instituições; a pulverização das linguagens algorítmicas e estatísticas; as exigências crescentes por transparência, precisão e celeridade; a vinculação entre o tratamento de dados e as possibilidades de desenvolvimento democrático e sustentável. É dada especial atenção às transformações éticas e epistemológicas que envolvem os *big data*, bem como os efeitos da mobilização de conhecimentos computacionais por especialistas, incluídos aí os/as próprios/as antropólogos/as, em diferentes ambientes de produção de conhecimento.

**Palavras-chave:** Antropologia; *big data*; Dados; Comparação; Tecnologias digitais.

**BIG DATA: WAYS OF DOING,  
COMPARING AND GOVERNING****Abstract**

In this dossier, we bring together researches that describes and analyzes relationships under the effect of *big data*. From sites that are quite different from each other — the legal, the bureaucratic, the agricultural, the biotechnological, the economic, the pornographic — we turn to the conflicting and creative agencies involved in what *data* makes us see and what they are capable of creating through their potential for hyper-relationality. We discuss the relationship between voluminous databases and various devices, such as applications, satellites, microscopes, documents and algorithms, capable of arranging and distributing informational elements through comparative practices. We examine *big data* collection and processing procedures in different sectors and institutions; the spread of algorithmic and statistical languages; the growing demands for transparency, precision and speed; and the link between data processing and the possibilities of democratic and sustainable development. Special attention is paid to the ethical and epistemological transformations involving *big data*, as well as the effects of the mobilization of computational knowledge by experts, including anthropologists themselves, in different knowledge production environments.

**Keywords:** Anthropology; *big data*; Data; Comparison; Digital technologies.

## **BIG DATA: MODOS DE HACER, COMPARAR Y GOBERNAR**

### **Resumen**

En este dossier reunimos investigaciones que describen y analizan las relaciones creadas bajo el efecto del *big data*. Desde sitios muy diferentes — el legal, el burocrático, el agrícola, el biotecnológico, el económico, el pornográfico — pasamos a las agencias conflictivas y creativas involucradas en lo que los *datos* nos hacen ver y lo que son capaces de crear a través de su potencial para hiperrelacionalidad. Luego discutimos la relación entre voluminosas bases de datos y diversos dispositivos, como aplicaciones, satélites, microscopios, documentos y algoritmos, capaces de ordenar y distribuir elementos informativos a través de prácticas comparadas. Examinamos los procedimientos de recopilación y procesamiento de *big data* en diferentes sectores e instituciones; la difusión de lenguajes algorítmicos y estadísticos; las crecientes exigencias de transparencia, precisión y rapidez; el vínculo entre el procesamiento de datos y las posibilidades de un desarrollo democrático y sostenible. Se presta especial atención a las transformaciones éticas y epistemológicas que involucran *big data*, así como a los efectos de la movilización de conocimiento computacional por parte de expertos, incluidos los propios antropólogos/as, en diferentes entornos de producción de conocimiento.

**Palabras clave:** Antropología; *big data*; Datos; Comparación; Tecnologías digitales.



**Sara R. Munhoz** é mestra e doutora pelo PPGAS/UFSCar. Atualmente, realiza estágio pós-doutoral no Departamento de Política Científica e Tecnológica da Unicamp (Processo Fapesp 2023/13441-7). Pesquisadora associada ao GEICT (Unicamp), ao Hybris (USP/UFSCar) e ao Núcleo de Antropologia da Política (Museu Nacional). Autora dos livros *O Governo dos Meninos: liberdade tutelada e medidas socioeducativas* (Edufscar, 2017) e *A paixão do acesso: uma etnografia das ferramentas digitais e da jurisprudência do STJ* (no prelo, Editora Hucitec). Tem experiência em antropologia política, dos documentos, da família, do Estado, do direito e dos dados.

**Vanessa Parreira Perin** é doutora em Antropologia Cultural pelo Programa de Pós-Graduação em Sociologia e Antropologia da Universidade Federal do Rio de Janeiro. Durante o doutorado realizou pesquisa sobre a cooperação internacional brasileira em Moçambique, na área de desenvolvimento agrícola e transferência de tecnologias em agricultura tropical. Atualmente, é pesquisadora de pós-doutorado no PPGAS da UFSCar, desenvolvendo pesquisa sobre a cadeia de suprimentos do complexo soja e seus mecanismos tecno-financeiros para a contínua expansão da fronteira agrícola através do agronegócio.

**Magda dos Santos Ribeiro** é professora do Departamento de Antropologia e Arqueologia e do Programa de Pós-Graduação em Antropologia da Universidade Federal de Minas Gerais. É pesquisadora associada ao Laboratório de Antropologia das Controvérsias Sociotécnicas – LACS/UFMG. Doutora (2016) e mestre (2011) em Antropologia Social pela Universidade de São Paulo com período sanduíche na Université Sorbonne Nouvelle, Paris III (2012-2013). Foi Rutherford Fellow na Universidade de Manchester/UK (2018-2019). Realiza pesquisa de campo com povos castanheiros na região do Jari, Amapá, Amazônia oriental. Suas pesquisas buscam compreender as relações entre conhecimentos tradicionais e práticas tecnocientíficas por meio de uma perspectiva comparada, com ênfase nos seguintes temas: economias da troca; cultura material e consumo; meio ambiente e capitalismo.

---

**Editora-Chefe:** María Elvira Díaz Benítez

**Editor Adjunto:** John Comeford

**Editor Associado:** Luiz Costa

**Recebido em:** 12 Março 2023

**Aceito em:** 9 Julho 2024